

CLAIMS

What is claimed is:

1. A speaker authentication system, comprising:
 - an input receptive of user speech from a user;
 - an extraction module adapted to extract acoustic correlates of aspects of the user's physiology from the user speech, including at least one of glottal source parameters, formant related parameters, timing characteristics, and pitch related qualities; and
 - an output communicating the acoustic correlates to an authentication module adapted to authenticate the user by comparing the acoustic correlates to predefined acoustic correlates in a datastore.
2. The system of claim 1, wherein said extraction module is adapted to extract glottal source parameters that include spectral qualities.
3. The system of claim 1, wherein said extraction module is adapted to extract glottal source parameters that include breathiness.
4. The system of claim 1, wherein said extraction module is adapted to extract glottal source parameters that include noise content.

5. The system of claim 1, wherein said extraction module is adapted to extract glottal source parameters that include at least one of jitter and shimmer related to fluctuations in pitch period and amplitude.

6. The system of claim 1, wherein said extraction module is adapted to extract glottal source parameters that include glottal source waveform shape related to phase information.

7. The system of claim 1, wherein said extraction module is adapted to extract formant related parameters that include a pattern of high formants related to head shapes and cavities.

8. The system of claim 1, wherein said extraction module is adapted to extract formant related parameters that include an estimate of vocal tract length.

9. The system of claim 1, wherein said extraction module is adapted to extract formant related parameters that include low formant patterns related to at least one of accent and dialect.

10. The system of claim 1, wherein said extraction module is adapted to extract formant related parameters that include an estimate of nasality related to velum opening.

11. The system of claim 1, wherein said extraction module is adapted to extract formant related parameters that include formant bandwidth.

12. The system of claim 1, wherein said extraction module is adapted to extract timing characteristics at a phoneme level.

13. The system of claim 1, wherein said extraction module is adapted to extract pitch related qualities that include characteristics derived from clustered training data.

14. The system of claim 1, further comprising a dialogue manager adapted to require the user to copy speech of a prompt when providing the user speech.

15. The system of claim 1, further comprising a dialogue manager adapted to require the user to perform a distracting task while providing the user speech input.

16. The system of claim 1, further comprising a scoring mechanism adapted to combine multiple feature sets differentiated according to modality using combining weights that are sensitive to changes in context and environment.

17. The system of claim 1, further comprising a communications network conveying the acoustic correlates to the authentication module, wherein the authentication module is adapted to generate an authentication decision and transmit the decision across the network to an input of the speaker authentication system.

18. A speaker authentication method, comprising:

- receiving user speech from a user;
- extracting acoustic correlates of aspects of the user's physiology from the user speech, including at least one of glottal source parameters, formant related parameters, timing characteristics, and pitch related qualities; and
- communicating the acoustic correlates to an authentication module adapted to authenticate the user by comparing the acoustic correlates to predefined acoustic correlates in a datastore.

19. The method of claim 18, further comprising extracting glottal source parameters that include spectral qualities.

20. The method of claim 18, further comprising extracting glottal source parameters that include breathiness.

21. The method of claim 18, further comprising extracting glottal source parameters that include noise content.

22. The method of claim 18, further comprising extracting glottal source parameters that include at least one of jitter and shimmer related to fluctuations in pitch period and amplitude.

23. The method of claim 18, further comprising extracting glottal source parameters that include glottal source waveform shape related to phase information.

24. The method of claim 18, further comprising extracting formant related parameters that include a pattern of high formants related to head shapes and cavities.

25. The method of claim 18, further comprising extracting formant related parameters that include an estimate of vocal tract length.

26. The method of claim 18, further comprising extracting formant related parameters that include low formant patterns related to at least one of accent and dialect.

27. The method of claim 18, further comprising extracting formant related parameters that include an estimate of nasality related to velum opening.

28. The method of claim 18, further comprising extracting formant related parameters that include formant bandwidth.

29. The method of claim 18, further comprising extracting timing characteristics at a phoneme level.

30. The method of claim 18, further comprising extracting pitch related qualities that include characteristics derived from clustered training data.

31. The method of claim 18, further comprising requiring the user to copy speech of a prompt when providing the user speech.

32. The method of claim 18, further comprising requiring the user to perform a distracting task while providing the user speech input.

33. The method of claim 18, further comprising combining multiple feature sets differentiated according to modality by using combining weights that are sensitive to changes in context and environment.

34. The method of claim 18, further comprising:
conveying the acoustic correlates to the authentication module via
a communications network; and
receiving an authentication decision generated by the
authentication system via the communications network.